DOCUMENT RESUME

ED 174 004 FL 010 291

AUTHOR Miner, Kenneth L.
TITLE On the Notion "Restricted Linguistic Theory": Toward Error-free Data in Linguistics. Kansas Working Papers in Linguistics, Volume 3, 1978.
INSTITUTION Kansas Univ., Lawrence. Dept. of Linguistics.
PUB DATE 78
NOTE 20p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Contrastive Linguistics; Data Analysis; Deductive Methods; *Descriptive Linguistics; Diachronic Linguistics; *Language Research; *Language Universals; *Linguistics; *Linguistic Theory; Researchers; Research Methodology; *Scientific Methodology; Structural Linguistics
IDENTIFIERS Language Families

ABSTRACT

The construction of general linguistic theories is subject to errors of fact and perspective. The theoretical linguist is inclined by the nature of his work to fall into such errors because he/she is obliged to generalize about many languages on the basis of research done in one or a few. The scope of this task inevitably leads to a failure to consider all facts in their context and renders the completion of all necessary background research impossible. C. Douglas Johnson's article "Global Constrained Rule Application in Menomini" is analyzed to support this position. By limiting themselves to restricted linguistic theories involving a specific family of languages, theoretical linguists will avoid inevitable error and remain true to the nature of their field, since experience has shown that only within language familes do interesting implicational universals seem to hold. A brief bibliography is appended. (JB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
U. of Kansas — Patricia J. Hamel
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED174004

RL010291

ON THE NOTION "RESTRICTED LINGUISTIC THEORY": Toward Error-Free Data in Linguistics

Kenneth L. Miner

> There is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of fact.
> -Mark Twain

Abstract

This paper has to do with the dichotomy generalism *vs.* particularism in linguistics; that is, the distinction between "theoretical linguistics" and the actual study of languages.[1] I will urge (a) that theoretical linguistics is generally working with data containing errors of fact or of perspective, (b) that this is not necessarily due to mere carelessness but inheres in the practice of theoretical linguistics itself, (c) that this fact overrides normal principles of theory reduction which otherwise would prohibit restricted linguistic theories[2] in favor of general linguistic theories, (d) that general linguistic theories *must* be replaced by restricted linguistic theories if linguistic theories are to be constructed on the basis of error-free data, and (e) that the effect of this conclusion on the field ought to be to remove "theoretical linguistics"[3] as an activity independent of the actual study of languages.

Errors of Fact and of Perspective

Descriptions of languages contain information; this information is processed into data for theory construction. It is during this processing that errors develop. We may categorize such errors into three types:

(i) simple errors of fact, presumably due to carelessness;
(ii) subtle errors of fact, due to lack of *thorough* study of linguistic descriptions;
(iii) errors of perspective, due to *normalization of variables* in linguistic descriptions.

Beyond dutifully deploring their occurrence, I have little to say about simple errors of fact. An example would be Lightner's (1971:521) statement that Winnebago, a Siouan language spoken in Wisconsin and Nebraska, has final obstruent *voicing*, apparently due to confusion of Winnebago with Dakota combined with lack of correct information on Dakota. Much more interesting are subtle errors of fact and errors of perspective; although difficult and rather time-consuming to exemplify,

these result (I maintain) from unavoidable limitations upon the time a theoretician is able to spend with descriptions of individual languages. I will give examples of these types of error in the following section, then return to the general discussion.

Examples

Johnson 1972 offers a supposed argument from Menominee (a Central Algonquian language spoken near Green Bay, Wisconsin) for a global negative constraint in the phonology of that language, thus supporting one side of the then especially prominent issue whether phonological theory should allow the application of rules to be globally constrained.

In order to prepare for his argument, Johnson first revises Bloomfield's (1939, 1962) ordering of three important rules governing predictable vowel length, rules which Johnson calls Checked/Free Vowel Shortening, Even Vowel Adjustment, and Second-Vowel Lengthening.[4] (The reader who is already familiar with Bloomfield's classic works on this language, or who is otherwise familiar with the processes in question, may simply skip what follows and proceed directly to the heading Johnson's Ordering Argument.)

Checked/Free Vowel Shortening (CFVS) states that a vowel is short after a cluster if not before a cluster (*i.e.*, short after a cluster in an open syllable, although in Menominee "open syllable" includes (C)VC#).

$$\text{CFVS:}\quad V \rightarrow \breve{V} \;/\; CC __ C\{\substack{\#\\V}\}$$

Consider the inflectional ending realized as either -ɛw or -ɛ•w, added to a verb stem to indicate a third person acting upon a fourth person. If the stem ends in a cluster, the vowel of this ending is short; otherwise, it is long:

me•n-ɛ•w	'he gives it to him'
ko•hn-ɛw	'he swallows him'
nɛ•w-ɛ•w	'he sees him'
nɛqn-ɛw	'he kills him'
ni•ma•h-ɛ•w	'he gives him permission'
meʧɛ•hn-ɛw	'he tracks him'

Between clusters, however, a vowel does not shorten. Consider the roots ke•hka-/ke•hka•- 'berate', pa•qta-/pa•qta•- 'catch; betray', and a•gɛ-/a•qtɛ•- 'be extinguished'. In these roots, the final vowels are preceded by clusters. If not followed by clusters in actual words, these vowels are short:

ke•hkamɛw	'he berates him'
pa•qtamɛw	'he tells on him'
a•qtɛw	'it goes out'

But if the root-final vowels occur before a cluster, and thus are between clusters, they appear as long:

ke•hka•htam	'he berates it'
pa•qta•hsen	'he is stuck, impeded'
a•qtɛ•qnen	'it is extinguished by the wind'

While it is always the case in Menominee that vowels are short after clusters in open syllables (except in the second syllables of non-glottal words; see below), it is not the case that vowels are always long in the complementary environment; that is, between clusters or after a single consonant:

kaka•qtohnɛw	'he barely hobbles along'
mi•nekɛ•hsɛhkɛw	'he makes a path'
nɛhe•qtohnɛw	'he makes him adept'
pɛ•ckaqtaw	'he puts it carefully in place'

Therefore the vowels which alternate in these morphemes are underlyingly long, and are shortened in the proper environment by CFVS.

Even-Vowel Adjustment (EVA) states that in an even-numbered syllable, vowels are long before clusters and short before single consonants (*i.e.*, long in closed syllables and short in open syllables), where "even-numbered syllables" are determined by counting from the next preceding long vowel in non-glottal words, and by counting from the beginning of the word in the case of glottal words, where "glottal words" are words having short initial syllables terminating in q (glottal stop).

$$\text{EVA:}\quad V \rightarrow \{\overset{\breve{V}}{\bar{V}} / _ CC\} \quad / \quad \{\overset{<\#>}{\bar{V}}\}C_o\breve{V}<\text{glottal}>_oC_o(VC_o\breve{V}C_o)^* __$$

This rule gives rise to extensive alternations in the shapes of surface morphemes.[5] Consider the following:

se•wa•poh	'vinegar'
esko•tɛ•wa•poh	'whiskey'

The above words contain the final element -apw/-a•pw 'liquid' (followed by a non-significant - in this case - ending /-ɛh/ which always follows the element -apw/-a•pw; a rule not otherwise relevant to this discussion converts /wɛ/ to o after a consonant; thus /-apw-ɛh/ → -apoh, etc.). The element -apw/-a pw always has a long vowel in positions other than in an even-numbered syllable, as in the words just given. But in an even-numbered syllable, the vowel is short:

maski•hkiwapoh	'tea'
keme•wanapoh	'rain water'
aske•tɛ•menapoh	'lemonade'
we•yakapoh	'various drinks'

So much for the shortening effect of EVA. It's lengthening effect can be exemplified by word-final clusters, which are reduced to their first member unless a suffix, for example, a plural suffix, is added. Note the lengthening

of vowels in even-numbered syllables before clusters in the following pairs of singular and plural nouns.

a•sepɛh	'rock; cliff'
a•sepɛ•hkok	'(plural)'
ka•ka•kemenakɛh	'staghorn sumac'
ka•ka•kemenakɛ•hsyan	'(plural)'
tu•kaniq	'blackbird'
tu•kani•qsak	'(plural)'

Finally, Second-Vowel Lengthening (SVL) lengthens the second vowel of non-glottal words if the first vowel is short:

$$\text{SVL:}\quad V \rightarrow \bar{V} \;/\; \#C_o\breve{V}[\text{-glottal}]_o __$$

In non-glottal words, it is regularly the case that one (or both) of the first two vowels is long:[6]

sa•konɛnehtam	'he remembers it'
ne•hnawat	'the weather relents'
tepa•hekan	'watch; hour; mile'
wasc•hsyan	'water weed'
wɛ•wɛ•nen	'Thank you'
ka•ka•kemenakɛh	'staghorn sumac'

Note, however, the alternation in vowel length depending upon whether or not a prefix is present. Consider the following words:

mase•nahekan	'book; paper'
sakɛ•hkam	'he holds it with his foot'
mace•kew	'he is in bad shape'

In the above words, the second vowels are underlyingly short and are lengthened by SVL, as can be seen by their prefixed forms, in which different vowels come to be second in the word:

nema•senahekan	'my book; paper'
nesa•hkɛhkan	'I hold it with my foot'
nema•cekem	'I am in bad shape'

Of course there are stems having underlyingly long first vowels; these undergo no change when prefixed:

me•kon	'feather'
neme•konem	'my feather'
me•san	'he sews it'
neme•sa•n	'I sew it'

Glottal words, as we have said, behave exceptionally with respect to EVA and SVL. Note the length alternation in the -ɛw/-ɛ·w suffix we have discussed, when it is added to stems having short initial syllables terminating in glottal stop (q):

aqsekanɛw	'he picks him up'
peqtɛnɛ·w	'he touches him by mistake'

As noted earlier, in the case of glottal words one counts from the beginning of the word in determining which are the even-numbered syllables for the purposes of EVA, rather than only from the nearest preceding long vowel. Of course long vowels may occur in glottal words, in which case one counts from them, but in such words there is no *guaranteeable* long vowel (as there always is in the case of non-glottal words), precisely because glottal words do not undergo Second-Vowel Lengthening:

naqnɛh	'mitten'
paqnakow	'he drowns'
sɛqsiwahekan	'whip'
mɛqnapɛ·w	'giant'
soqsoma·kɛ·w	'he slides; coasts'
koqtam	'he fears it'

Of course, if glottal words are prefixed, then SVL applies to lengthen their second syllables:

nena·qnɛh	'my mitten'
nepa·qnakim	'I drown'
nesɛ·qsiwahekan	'my whip'
neso·qsomakɛ·m	'I slide; coast'
neko·qtan	'I fear it'

Johnson's Ordering Argument

We first take up Johnson's ordering argument. While it contains no error of fact, it does contain an error of perspective and will serve to illustrate this type of error. The error consists in considering only data which happened to be relevant to Johnson's momentary inquiry, in a case where, had other data been considered, the reasoning behind Bloomfield's decisions would have been obvious.

Bloomfield, in his famous 'Menomini morphophonemics' (Bloomfield 1939), set up the ordering CFVS-SVL-EVA. Johnson notes that this ordering is possible only if EVA is stated as we have stated it on p. 3: even-numbered syllables are determined by counting from the nearest preceding long vowel in the case of non-glottal words, and from the beginning of the word (only) in the case of glottal words. This is essentially the way the rule is stated in Bloomfield's 1939 paper. Johnson observes that this formulation is unnecessarily complicated, and is complicated only to prevent EVA from incorrectly shortening a second syllable that is underlyingly long or that has been lengthened by SVL: if EVA were simply stated so as to count even-

numbered syllables *in any sequence of short syllables*, there would be no need to treat glottal words exceptionally. What Johnson has in mind is a rightward-iterative counting rule of the type:

$$V \rightarrow \text{even-numbered} \;/\; \breve{V}C_0 __ \qquad \text{(rightward iterative)}$$

(The entire rule cannot be formalized iteratively at the present stage of iterative rule theory; for some discussion, see Howard 1972:§7.2 and Miner 1975:Ch. 2.) If this simplification were made, the ordering of Bloomfield 1939 would of course not work:

given:	/natomenan/	/kaki·pɛhtɛw/
CFVS	-	-
SVL	nato·menan	-
EVA	natomenan	kakipɛhtɛw
result:	*natomenan	*kakipɛhtɛw

(I have added the second derivation to make Johnson's case clearer; the glosses are 'when I call thee' and 'he is deaf'.)

Now in Bloomfield 1962 (rather than 1939) the ordering of our length alternation rules is given as (or at least the rules are presented in the order) SVL-CFVS-EVA; that is, in the 1962 work SVL applies first. This ordering, for Johnson, is even worse than that of Bloomfield 1939, since now not only EVA but also CFVS must be complicated so as not to incorrectly shorten second syllables which are underlyingly long or which have been lengthened by SVL. If CFVS were simply as we have stated it on p. 2, that is,

$$V \rightarrow \breve{V} \;/\; CC __ C\{{}^{\#}_{V}\}$$

then derivations in the 1962 ordering would be incorrect:

given:	/pahnɛ·w/	/kohkapew/
SVL	-	kohka·pew
CFVS	pahnɛw	kohkapew
EVA	-	-
result:	*pahnɛw	*kohkapew

To avoid this, Bloomfield 1962 gives CFVS as follows (§4.51):

> *Anywhere after the first long vowel of a non-glottal word, and anywhere at all in a glottal word,* a long vowel preceded by a consonant cluster and not followed by a consonant cluster is shortened.

(Emphasis mine. I have added the second derivation above to make Johnson's case clearer; the words mean 'he roasts him in the fire' and ' he turns in his seat'.)

Noting that in Bloomfield 1962 *all three rules* have to mention special treatment of glottal words, Johnson advocates simplifying EVA and CFVS as shown above, and applying the rules in the order CFVS-EVA-SVL. If this is done, the only rule that behaves exceptionaly with respect to glottal words is SVL: the second vowel of glottal words is not lengthened by this rule. Johnson's derivations go as follows:

given:	/natomenan/	/kaki·pɛhtɛw/	/pahnɛ·w/	/kohkapew/
CFVS	-	-	pahnɛw	-
EVA	-	-	-	-
SVL	nato·menan	-	pahnɛ·w	kohka·pew
result:	nato·menan	kaki·pɛhtɛw	pahnɛ·w	kohka·pew

Johnson's rule simplifications and reordering of the rules give the correct results for glottal words as well:

given:	/soqsoma·kɛ·w/	/ne-soqsoma·kɛ·m/
CFVS	-	-
EVA	-	nesoqsomakɛ·m
SVL	-	neso·qsomakɛ·m
result:	soqsoma·kɛ·w	neso·qsomakɛ·m

Johnson remarks at this point that Bloomfield's ordering "would be rejected by the orthodox principles of current generative phonology". This I find rather misleading. Note Hockett's editorial remark (Bloomfield 1962: viii):

> It is quite clear that what Bloomfield really wanted to accomplish, in his description of the language, was something matching in completeness and accuracy Panini's grammar of Sanskrit, for which he had such great respect.

Not stopping at mere admiration of Panini, Bloomfield was one of the founders of modern process phonology and well understood simplicity of description (his term: convenience) insofar as rule statement and ordering of rules are concerned. It may be true (or may have been true in 1972) that his solution violates "currently accepted cannons [sic] of linguistic analysis" (Johnson appears to have in mind the simplicity metric, by which he himself must have been guided in this paper), but there have really been no changes in the *principles* of process description since Bloomfield's time.

Why, then, did Bloomfield not "see the obvious" - why did he not simplify his rules and order them as Johnson now advocates?

The answer becomes apparent as soon as we go beyond Johnson's examples and apply his ordering to other types of input. We give below Johnsonian derivations for (a) words which have underlyingly long second vowels not followed by clusters, e.g., kaki·pɛhtɛw 'he is deaf', (b) prefixed forms of (a), e.g., naka·ki·pɛhtɛm 'I am deaf'; (c) words which have underlyingly long second vowels followed by clusters, e.g., tata·skikatɛ·w 'he is short-legged'; (d) prefixed forms of (c), e.g., nata·ta·skikatɛ·m 'I am short-legged; (e) words which have underlyingly short first and second syllables

with the second syllables terminating in clusters, e.g., sakɛ•hkam 'he holds it with his foot'; (f) prefixed forms of (e) e.g., nesa•kɛhkan 'I hold it with my foot'; (g) words which have underlyingly short second vowels preceded but not followed by clusters, e.g., kohka•pew 'he turns in his seat'; (h) prefixed forms of (g), e.g., neko•hkapem "I turn in my seat'; (i) words which have underlyingly long second vowels preceded but not followed by clusters, e.g., wackɛ•ta•cekow 'he goes by detour'; and (j) prefixed forms of (i), e.g., newa•ckɛtacekim 'I go by detour'.

Johnsonian Derivations (order CFVS-EVA-SVL)

given:	/kaki•pɛhtɛw/	/ne-kaki•pɛhtɛm/	/tata•skikatɛ•w/	/ne-tata•skikatɛ•m/
CFVS	-	-	-	-
EVA	kakipɛhtɛw	-	-	-
SVL	kaki•pɛhtɛw	neka•ki•pɛhtɛm	-	neta•ta•skikatɛ•m
result:	kaki•pɛhtɛw	neka•ki•pɛhtɛm	tata•skikatɛ•w	neta•ta•skikatɛ•m

given:	/sakɛhkam/	/ne-sakɛhkan/	/kohkapew/	/ne-kohkapem/
CFVS	-	-	-	-
EVA	sakɛ•hkam	-	-	neko•hkapem
SVL	-	nesa•kɛhkan	kohka•pew	-
result:	sakɛ•hkam	nesa•kɛhkam	kohka•pew	neko•hkapem

given:	/wackɛ•ta•cekow/	/ne-wackɛ•ta•cekim/
CFVS	wackɛta•cekow	newackɛta•cekim
EVA	-	newackɛtacekim
SVL	wackɛ•ta•cekow	newa•ckɛtacekim
result:	wackɛ•ta•cekow	newa•ckɛtacekim

Now there is no question that these derivations give the correct outputs. But such an analysis would never have been acceptable to Bloomfield. Note, for example, that in the derivation of kaki•pɛhtɛw the second vowel appears to have been lengthened by rule, i.e., appears to be predictably long, which is contrary to fact, since it is underlyingly long. But in the derivation of the prefixed form of this same item, neka•ki•pɛhtɛm, this same vowel is *not* lengthened by rule, i.e., is unpredictably long. Thus for stems of this underlying shape, their second vowels are portrayed as predictably long in unprefixed forms and unpredictably long in prefixed forms. What would Bloomfield have set up as the base forms of these stems? Further note that in the case of tata•skikatɛ•w, neta•ta•skikatɛ•m the second vowel of the stem is unpredictably long in both the prefixed and the unprefixed form. In the derivation of sakɛ•hkam the second vowel is lengthened by EVA, not by SVL, while in the derivation of nesa•kɛhkan the second vowel is lengthened by SVL. Thus for items of this underlying shape, the second vowel is lengthened by EVA if unprefixed, by SVL if prefixed. The converse holds for items like kohka•pew, neko•hkapem.

In Bloomfieldian derivations for all of these items (and for all the words of the language), all of the long second vowels are predictably lengthened if and only if not underlyingly long, and all such are lengthened by the same rule, SVL:

Bloomfieldian Derivations (order SVL-CFVS-EVA)

given:	/kaki·pɛhtɛw/	/ne-kaki·pɛhtɛm/	/tata·skikatɛ·w/	/ne-tata·skikatɛ·m/
SVL	-	neka·ki·pɛhtɛm	-	neta·ta·skikatɛ·m/
CFVS	-	-	-	-
EVA	-	-	-	-
result:	kaki·pɛhtɛw	neka·ki·pɛhtɛm	tata·skikatɛ·w/	neta·ta·skikatɛ·m

given:	/sakɛhkam/	/ne-sakɛhkan/	/kohkapew/	/ne-kohkapem/
SVL	sakɛ·hkam	nesa·kɛhkan	kohka·pew	neko·hkapem
CFVS	-	-	-	-
EVA	-	-	-	-
result:	sakɛ.hkam	nesa·kɛhkan	kohka·pew	neko·hkapem

given:	/wackɛ·ta·cekow/	/ne-wackɛ·ta·
SVL	-	newa·ckɛ·ta·cekim
CFVS	-	newa·ckɛta·cekim
EVA	-	newa·ckɛtacekim
result:	wackɛ·ta·cekow	newa·ckɛtacekim

Now of course we do not have direct access to Bloomfield's reasoning, but since he presumably saw process description as a convenient way of acacounting, in a straightforward and coherent manner, for predictable alternations, he surely would have rejected a Johnsonian analysis as wantonly formalistic and - most importantly - anything but simple.

As for learnability of the rules, a modern theoretician following Kiparsky[7] would note the considerable opacity of the Johnsonian solution: according to the definition of Kiparsky 1973, a process of the form A → B/C_D is opaque to the extent that there are phonetic forms in the language having

i. A in env. C_D
ii. B derived by the process P in env. other than C_D
iii. B not derived by the process P (i.e., underlying or derived by another process in env. C_D

Given Johnson's maximally simple rule formulations, CFVS is opaque with respect to to·kohka·pew and wackɛ·ta·cekow by (i), and EVA is opaque with respect to kaki·pɛhtɛw, neka·ki·pɛhtɛm, neta·ta·skikatɛ·m, nesa·kɛhkam, kohka·pew, and wackɛ·ta·cekow, all by (i); and SVL is opaque with respect to sakɛ·hkam and neko·hkapem by (iii).

Thus Bloomfield's reasoning seems clear: all rules state true generalizations about the surface, and each rule makes a separate generalization; there are no "false starts" in derivations. Or in modern terms, Bloomfield's derivations show no opacity whatever (while in Johnson's, although the rules are simpler *in statement*, every one of them is opaque). Thus Bloomfield's revision of the order of application of the length alternation rules in his 1962 work, as over the 1939 paper, was an *improvement* in his own terms, since in the 1939 paper these goals were violated by his statement of CFVS.[8]

Johnson's Global Reference Argument

We turn now to the major topic of Johnson's paper, the claim that Menominee phonology requires global reference in one of its rules.

In preparation for this discussion we need to mention two more rules of Menominee phonology, Epenthesis and Raising.

Epenthesis inserts e between sequences, arising in derivation, of non-syllabic and consonantal segments: /po•n-m-ɛ•w/ → po•nemɛw 'he stops talking to him', etc.

Epenthesis: $\emptyset \rightarrow e \;/\; [-syll] + __ \;[+cons]$

Raising raises the long mid vowels e• and o• respectively to i• and u• when they are followed later in the stem by a high vowel or post-consonantal glide w, y, unless a low front vowel ɛ, ɛ• intervenes. (This rule should apply iteratively, but like Even Vowel Adjustment cannot be completely formalized at the present stage of iterative rule theory; the formulation below is a makeshift):

Raising: $\begin{vmatrix} e\bullet \\ o\bullet \end{vmatrix} \rightarrow \begin{vmatrix} i\bullet \\ u\bullet \end{vmatrix} \;/\; __ \; X \; \{^{Cw,\ Cy}_{i,\ u}\}$

Condition: X does not contain a low front vowel, word boundary, or preverb boundary.

Thus:

/ke•ketow/ → ke•ketow	'he speaks'
/neke•ketim/ → neki•ketim	'I speak'
/so•pomah/ → so•pomah	'maple sugar'
/so•pomahkwan/ → su•pomahkwan	'maple sugar making'

Johnson first discusses cases of Regular Haplology in Menominee.[9] Certain suffixes must be lexically marked to undergo this minor rule; among them are -owa•w (pluralizer of non-1 p possessor) and -nɛme- 'be overtaken by time or weather'; these are reduced to -a•w and -ɛme- respectively, when following a morpheme ending with w and n respectively. Note the derivations of 'their tails' and 'he is caught in the rain':

given:	/oso•w-owa•w-an/	/kemewan-nɛme-w/
Epen.	-	kemewanenɛmew
Hapl.	oso•wa•wan	kemɔwanɛmew
length rules	-	keme•wanɛmew
result:	oso•wa•wan	keme•wanɛmew

Johnson formally states Regular Haplology as follows (which suffices for our present purposes):

Regular Haplology:

CVC' becomes C if C is a non-syllabic, V is a vowel,
C' is identical to C in phonological features, and
C' is specified as #H.

The #H is simply a diacritic marking certain morphemes as undergoing the rule. Setting up the ordering Epenthesis-Regular Haplology-CFVS-EVA-SVL-Raising, Johnson then notes an apparent difficulty in deriving the word tahki•kamiw 'it is cool water' (root tahk- 'cool; cold'; -kamy- #H 'water; liquid'; -w (inflectional)).[10] The derivation, says Johnson, ought to go as follows:

given:	/tahk-kami• w/
	#H
Epen.	tahkekami•w
Hapl.	tahkami•w
CFVS	-
EVA	-
SVL	tahka•mi•w
Raising	-
result:	*tahka•mi•w

Since the above ordering of rules is proper for all other derivations, Johnson concludes that Regular Haplology does not apply in this case because Regular Haplology is able to make a global reference to the output of Second Vowel Lengthening and "know beforehand" that the vowel it would normally delete is going to be long. This would allow the following derivation:

given:	/tahk-kami•-w/	
	#H	
Epen.	tahkekami•w	
Hapl.	-	(blocked by global reference)
CFVS	-	
EVA	tahkekamiw	
SVL	tahke•kamiw	
Raising	tahki•kamiw	
result:	tahki•kamiw	

The argument is sound, but not the data. Johnson unquestioningly assumes the second vowel in tahki•kamiw is epenthetic; however, it is not. It is a "premedial element" mentioned in Bloomfield 1962 at §19.14; the form at §15.15, tahki•hsewew 'it is a bit cold', where tahk-i•- is followed by diminutive -hs- and where an epenthetic e would not be in a raising environment, shows this same element. Since this surface i• is not in a raising environment, it cannot be epenthetic e. This -i•- shows up also

with other roots; instances are given at §19.14.

This is a very clear example of what I mean by a *subtle* error of fact: it is subtle because tahki•hsewew happens to be the *only* form in Bloomfield 1962 which indicates clearly (in spite of §19.14) that there are no grounds for taking the i• in tahki•kamiw to be epenthetic. I submit that, since a mere search for evidence relating to haplology would *not* lead one to §15.15, it is unreasonable merely to chide Johnson for "carelessness"; carelessness is not the problem. The problem is in the nature of theoretical linguistics: *there is no way to know beforehand what data are relevant, unless one studies languages in depth, which generalists by definition cannot do* As general linguistic theory is presently understood, there would be no hope of setting up any hypotheses of much interest if theoreticians limited themselves to just those languages they had time to learn.[11]

Johnson's other example for the global reference of Regular Haplology, kɛhke•ka•tɛ•p 'garter' is an old chestnut. It is inconsistent with its possessed form, nekɛ•hkatɛp 'my garter', the latter being haplologic, as Bloomfield mentions at §6.67 and §14.302, while the unpossessed form is not. Are there any other discrepancies between possessed and unpossessed forms in Menominee? Yes. But they do not show up in Johnson's inquiry. *Speakers* have trouble with "garter" too; one of my own informants in 1974 pronounced it four different ways, showing considerable confusion.

In addition, if Johnson had been able to consult all the materials for this language in the time available to him, he would have found that a discrepancy exists in Bloomfield 1962 with regard to this item, which is listed in Goddard, Hockett and Teeter 1972. The discrepancy is cleared up in Bloomfield 1975, but this was not in print when Johnson was writing (although it could have been checked in xeroxed form with just about anyone working on Central Algonquian in the U.S. or Canada); under the circumstances, only failure to consult (or even know about) Goddard, Hockett and Teeter 1972 prevented Johnson from removing this particular example from his paper as an unclear case.

It is remarkable in itself that Johnson bases his entire argument for global reference in Menominee on *two words*; eliminating one by appeal to unnoticed information in a linguistic description, we are left with one item as evidence, from this language, of the need for a theoretical innovation - an item which, for Johnson, should not have been used since it was unclear.

Errors Inevitable

It must be emphasized that the foregoing is *not* a mere polemic against Johnson. In fact, out of a number of theoreticians who have used Menominee "data" in recent years, Johnson is apparently the only one who bothered to use primary sources at all. But the very fact that Johnson could not have been expected to work his way through all of Bloomfield 1962, or consult with specialists, or (heaven forbid) do field work, merely in order to produce one fairly brief theoretical paper, constitutes the essential point I am attempting to make. As a theoretician, Johnson must obtain and process information from *many* languages and dialects and is not able to go into each one in depth.

If the issue were mere carelessness, we might, given a great deal of rather unappetizing and wasteful research, reconsider a famous counter of Chomsky & Halle (1965:138).

> ...Householder cites no evidence in substantiation of his charge that linguists who have been interested in or influenced by our work have no regard for linguistic fact and fail to meet common standards of accuracy and seriousness.

While I have always sympathized with Householder's insistence (often not made explicit enough) that even theoreticians should be scholars, I believe his charge to have been misdirected: errors in linguistic data are not attributable to mere carelessness, nor are they attributable, in my view, only to practitioners of only one theory or model; I believe rather that, given the task of theoretical linguistics as presently conceived, such errors are *inevitable*.

But, before we can maintain that errors are inevitable, we must argue that they are at least widespread, which will itself count as some evidence of their inevitability.

Generalizing on the Curve

One way to show that errors in linguistic data as used by theoreticians in theory construction are widespread would be to give a large number of examples. However, for any one individual, this would be difficult, since obviously no one individual has studied more than a few language families in depth. What we must do is rely on the observations of specialists, then "generalize on the curve".

Acceptance of this approach as a reasonable one depends partly upon one's experience *as* a specialist (particularist). Let us say, for example, that I specialize in some language family and that in the *majority* of cases in which, in the literature of general linguistic theory, data from this language family are cited, I find errors of fact or of perspective. Let us suppose further that I know two other specialists working in two other language families, and that they report the same finding. Now, going much further is difficult, because specialists who are willing to spend much time with the theoretical literature after repeated findings of this sort are rare, and for obvious reasons this is not the sort of information that gets into print. But it should be clear that at *some* point I must conclude that the problem is a general one. Generalizing on the curve is necessary simply because no one individual has the necessary knowledge to arrive singlehandedly at the conclusion. That does not mean that the conclusion cannot be arrived at.

Certain statistical frequencies (not in the formal sense, for I have not made counts) seem to support the notion that there is a general data problem in linguistics today. Perhaps the reader will concur with at least a subset of the following observations: (a) during discussion periods at theory conferences over half the issues raised have to do with primary *data*, rather than with theoretical claims, and yet (b) an error in someone's data

is generally treated as negligible in comparison to discovery of an error in someone's *reasoning*; (c) influential theory papers have been written on the basis of data from *one language* as described by *one researcher*; (d) theoretical discussions have been based on data from *one speaker*; (e) *second- and even third-hand data* are routinely tolerated in theoretical discussions; (f) information on languages well-covered in the literature and easily obtainable in any good university library is bypassed in favor of information from *personal acquaintances* of the theoretician (which means no account of the informants, the research process, etc.); (g) much theoretical discussion hangs upon a relatively small number of *examples* obtained often in the above ways; (h) under interrogation many a theoretical argument resolves to the form "if the data are right, then *p*", where *p* is a *hypothesis*; (i) claims that some phenomenon is frequent in "the world's languages" are routinely unaccompanied by *reference to research results* or figures of any kind.[12]

These sorts of observation make linguistics look very different from other sciences *even considering the times*, and if they are taken to be genuinely representative, theoretical linguistics can hardly even qualify as a scholarly discipline. It is embarrassing to note that in North American archaeology, for example, *most* of the really interesting hypotheses have to be rejected by disqualifying the original digs. One gets the impression that we do not feel linguistics to be a serious business.

Normalization

I have defined errors of perspective as involving *normalization*. What is being normalized is variables in the nature of linguistic descriptions. Johnson's error dealt with above was in failing to consider the possibility that Bloomfield may have had *reasons* for his particular analysis; in order to consider this possibility, Johnson would have had to look at phenomena outside the domain of his momentary inquiry; since he did not, the data *within* the domain of his inquiry has been normalized. Normalized data are, in a sense, data taken out of context. In this case, the result is the really rather unnerving conclusion that one of the founders of modern process description in phonology simply blundered in writing his rules in 1939 and then blundered again in the extended description which eventually surfaced in 1962.

Although it appears that all errors of perspective involve normalization of one sort or another, a complete typology will not be offered here. Such a typology should, however, attempt to characterize at least the following:

I. Errors of perspective on the scientific status of linguistics itself;
II. Errors of perspective on the integration of particular phenomena being treated within the language, language family, or language phylum as a whole;
III. Errors of perspective on normalizable variables in particular linguistic descriptions:
   A. training and experience of the author (quality and quantity);
   B. length and depth of exposure to the language;
   C. purpose of the description;
   D. theoretical interests/persuasions of the author;
   E. informants (number and quality);

F. author's techniques of data-gathering (translation? texts? etc.);
G. specialists' opinions of the description (reviews, etc.);
H. any evidence that the author *himself* normalized?

IV. Errors of perspective on the status of research on the language or on related languages;

V. Errors of perspective on socio- or psycholinguistic variables (multilingualism of informants; viability of the language; areal features; special effects of the methods of data-gathering; etc.)

VI. Errors in perspective on problems of analysis, especially if these are described in the literature and even more especially if they are described by the author himself.

Type I is prevalent, bordering at worst on pseudo-intellectualism (in this case, willingness to carry on academic activity without being able to state its intellectual credentials); similarly II; one now picks rules out of languages as if one were doing a study of comparative breechclouts (and with about the same results). III.D. has been exemplified here. IV leads the theoretician to treat as theoretical mysteries what are merely clear signs of insufficient pretheoretical investigation. V is straightforward, I believe.

Coming under VI (as well as V) is the fact that much more normalization than I have indicated so far is present in *all* theoretical discussions of Menominee. Much of it can be gotten from a *careful* reading of Bloomfield 1962: he had difficulty in distinguishing between -o·w- and -ow- (§1.23; cf. Miner 1977 for Goddard's example of this difficulty and its consequences); the lack of audible difference, for all practical purposes, between the two short vowels e and i is a very serious problem (§1.20; cf. Menomini Texts Bloomfield 1928, where the distinction isn't even written); there is a tendency to lengthen e in an even-numbered syllable even when it is *not* before a cluster (§1.21), suggesting that we need to look harder at the data before deciding on the length rules; Bloomfield found considerable speaker variation (remember that this was an obsolescent language even in the 1920s) and may *himself* have done some normalizing. All of these features of the sources for this language are *normalized out* in theoretical discussions.

Examples of second- and third-hand data are treated (inadequately, but at least treated) in Miner 1975.

Restricted Theories

The conviction that errors of fact or of perspective are unavoidable in general linguistic theories presses me toward the concept of "restricted" linguistic theories. An example of a restricted linguistic theory would be a theory of an individual language family - say, a theory of Algonquian languages, or a theory of Germanic languages, or even a theory of language L.[13]

Now a common assumption about linguistic theories is that they *must* be general; a theory of, say, Germanic languages would not be seen as having a natural domain of inquiry. By well-known principles of theory reduction, such a theory would, all else being equal, be subsumed under a general theory of language.

But it is precisely the burden of this paper that all else is *not* equal in this case. General linguistic theories and restricted linguistic theories utilize different sorts of data. Data for restricted linguistic

theories consists of facts; data for general linguistic theories consists of facts plus factual errors plus errors of perspective, these all deriving, I claim, from the method itself. The two kinds of theory are simply not the same, and the principles of theory reduction do not apply.[14]

It may be objected that there *are* no restricted linguistic theories to compare with general linguistic theories in the first place, to which I would respond with an *altes, wahres Wort*, namely, that there can be and often is a vast difference between what we think we are doing and what we are in fact doing. It is in fact the *general* theory that we do not have - it is rather within language families that interesting implicational universals seem to hold, that explanation seems to be within reach, that the prospect of making predictions (which *must* be part of explanation) seems real. This has been noticed by Lass 1972 as a general phenomenon and is implicit in most of the work on linguistic universals: research needs to be done to determine the exact extent to which work on general universals has actually unearthed family universals. Much of our practice clouds the issue: anyone who obtains his/her feel for what languages are like exclusively or even mainly from the theoretical literature is likely to have a radically different notion of what languages are like than particularists, which means a mistakenly high value placed upon even the most ephemeral "general hypothesis" and a corresponding underestimate of the usefulness of non-trivial family universals.

It may also be objected that the difference between general theories and restricted theories - the nature of their data - results from purely human limitations which are not relevant to the principles of theory reduction. But to maintain this position is to maintain the validity of pseudo-procedures - procedures which are admitted to be necessary but which are in point of fact impossible to carry out[15] - as part of the methodology of empirical science. It also reveals a serious misunderstanding of science: data *is and will always be* at the "human limitations" end of the relationship between science and the world, and, in spite of this, is the first and most crucial link in the methodological chain.

Once we see that restricted linguistic theories are not only scientifically "respectable" but are in fact "all we've got" at present, we can begin to formalize the concepts that underlie them, making explicit what has long been implicit in much linguistic work. Since theories constructed upon erroneous data are useless (assuming that we are serious), we have every reason to move toward theories constructed upon error-free data. This can be done if particularists and theoreticians are the same individuals and "theoretical linguistics" in the generalist sense is removed as an activity independent of the in-depth study of language.

Footnotes

1 I am of course aware that the distinction is not absolute; nonetheless, there are generalists who are only generalists and there are (even more) particularists who are only particularists.

2 Since seeing the typed version of Joseph Greenberg's 1977 LSA presidential address ('Rethinking linguistics diachronically', to appear, when revised,

in *Language*) I am embarrassed at not being sure that the notion is mine - since it was used by Greenberg and since I was present at his address. I *think* I used the expression in a discussion with someone the day before Greenberg's talk, but I honestly can't remember. It is not important; my discussion can only benefit by being seen as a footnote to Greenberg - but that is not to imply that he would agree with anything I say here, or (in particular) justify restricted theories the way I would.

3 My occasional quotes around this expression are intended, not to disparage "the real thing", but to convey annoyance at having to use the term "theoretical" to designate activities which often cannot even pretend to constitute science, hard or soft. It is my contention in this working paper, however, that even "the real thing" is in trouble.

4 I first give the rules as they are given in Bloomfield 1939.

5 Problems with the formalization of this process, especially as a directional iterative rule (see below), are discussed in Howard 1972 and Miner 1975.

6 There are small exception classes for each of the length alternation rules.

7 For Kiparsky's original formulation of opacity/transparency see Kiparsky 1971. I realize, of course, that for Kiparsky transparency is a tendency of grammars; for Bloomfield, we may say, it was a condition on them. (I refer of course not to his statements but to his practice.)

8 It should be noted at this point that, by apparent oversight, Bloomfield's 1962 statement of Even Vowel Adjustment is wrong; it is corrected in Goddard, Hockett and Teeter 1972 and in Bloomfield 1975.

9 The term is mine, not Bloomfield's or Johnson's. There is much *irregular* haplology in Menominee too; in fact if one looks at *all* of these tendencies in this language one will realize that here is ground on which angels fear to tread. Johnson certainly deserves credit for attempting to deal formally with part of the problem.

10 All underlying representations are simplified in this paper in order to conform with Johnson's assumptions, and for convenience.

11 I do not mean, of course, learn to speak. What I have in mind is the sense of "knowing a language" that Bloomfield referred to when he (reportedly) remarked, after reading an especially bad student paper, that "when comparing two languages, it helps to know one of them".

12 I intentionally refrain from citing examples; it is not the point of this paper to single out individuals for criticism, but to address a general problem in the field.

13 I mean explanatory theory, not mere grammatical description. I sidestep, for the moment, the question of the role of historical explanation as we

now understand it (or perhaps do not understand it) in restricted theories of language. I will not sidestep the issue of explanation, however: the goal of an empirical science is *explanation of events*. Hermeneutics and axiomatics have a role in linguistics, to be sure, but linguists *qua* linguists are likely to maintain their traditional interest in other people's languages as well as their own.

14 Note that differences in the nature of data are also precisely what keeps synchronic linguistics from automatically reducing to diachronic linguistics (on the ground that, if we had *complete* historical knowledge of a language family, we would not need synchronic explanations of any purely *linguistic* fact about any one of its members, and could assign any purely synchronic aspects of language organization or use to sciences other than linguistics). The comparison might be revealing if fully worked out: our data for historical linguistics are incomplete, just as our data for general linguistic theory, just as inevitably, contain errors. Therefore synchronic linguistics and restricted linguistic theory respectively are separate disciplines.

15 The notion "pseudo-procedure" is due to Abercrombie 1965.

References

Abercrombie, David, 1965. *Studies in Phonetics and Linguistics*. Oxford University Press (London).

Bloomfield, Leonard, 1928. *Menomini Texts. Publications of the American Ethnological Society* 12 (New York).

----------, 1939. Menomini morphophonemics *Travaux du Cercle Linguistique de Prague* 8.105-115. Repr. in Hockett, C.F., ed., *A Leonard Bloomfield Anthology*, Indiana University Press, 1970 (Bloomington and London).

----------, 1962. *The Menomini Language*. Edited by Charles F. Hockett. William Dwight Whitney Linguistic Series of Yale University. Yale University Press (New Haven and London).

----------, 1975. *Menominee Lexicon*. Edited by Charles F. Hockett. Milwaukee Public Museum Publications in Anthropology and History 3 (Milwaukee).

Chomsky, Noam and Morris Halle, 1965. Some controversial questions in phonological theory *Journal of Linguistics* 1.97-138. Repr. in Makkai, Valerie Becker, *Phonological Theory: Evolution and Current Practice*. Holt, Rinehart & Winston, 1972, 457-485.

Dingwall, William Orr, ed., 1971. *A Survey of Linguistic Science*. University of Maryland (College Park, Maryland).

Goddard, Ives, Charles F. Hockett & Karl Teeter, 1972. Some errata in Bloomfield's Menomini *IJAL* 38:1.

Howard, Irwin, 1972. *A Directional Theory of Rule Application in Phonology*. MIT PhD Dissertation. Mimeo, Indiana University Linguistics Club, 1973.

Johnson, C. Douglas, 1972. Globally constrained rule application in Menomini *Papers in Linguistics* 5:4.496-512.

Kiparsky, Paul, 1971. Historical Linguistics In Dingwall, ed. 1971.
----------, 1973. "Abstractness, opacity, and global rules" In O. Fujimura, ed., *Three Dimensions of Linguistic Theory*, TEC Corporation (Tokyo). Also Mimeo, Indiana University Linguistics Club, 1973. Also in Koutsoudas, A., ed. *The Application and Ordering of Grammatical Rules*, Mouton, 1975 (The Hague).
Lass, Roger, 1972. How intrinsic is content? Markedness, sound change, and 'family universals' *Edinburgh Working Papers in Linguistics* 1: 42-67.
Lightner, Theodore, 1971. Generative Phonology In Dingwall, ed., 1971.
Miner, Kenneth L., 1975. *Interference Relations in Menominee Phonology*. Indiana University PhD Dissertation
----------, 1977. Review of Bloomfield 1975. *IJAL* 43:1.66-73.